

# Application de l'analyse des séries chronologiques à la projection d'effectifs de population scolaire par la méthode des composantes

## Application of Time-Series Analysis to the Projection of School Enrollments by Cohort

Herbert L. Smith

Volume 38, numéro 1, printemps 2009

Enjeux de l'analyse démographique et nouvelles pistes méthodologiques

URI : <https://id.erudit.org/iderudit/039991ar>

DOI : <https://doi.org/10.7202/039991ar>

[Aller au sommaire du numéro](#)

Éditeur(s)

Association des démographes du Québec

ISSN

0380-1721 (imprimé)

1705-1495 (numérique)

[Découvrir la revue](#)

Citer cet article

Smith, H. L. (2009). Application de l'analyse des séries chronologiques à la projection d'effectifs de population scolaire par la méthode des composantes. *Cahiers québécois de démographie*, 38(1), 145-170.  
<https://doi.org/10.7202/039991ar>

Résumé de l'article

Cet article veut montrer qu'on peut réécrire des modèles démographiques en vue de réaliser des projections par cohorte, en les transposant dans un modèle économétrique vecteur autorégressif (VAR). De cette façon, la méthode des composantes se dote d'un cadre stochastique qui étend son envergure. Le potentiel de cette perspective est illustré à travers l'exemple d'une projection d'effectifs de population scolaire. Il met en valeur une série d'équations qui permet de vérifier la validité de plusieurs choix de modélisations habituellement utilisées dans le domaine de la prévision.

## **Application de l'analyse des séries chronologiques à la projection d'effectifs de population scolaire par la méthode des composantes**

**HERBERT L. SMITH\***

Cet article veut montrer qu'on peut réécrire des modèles démographiques en vue de réaliser des projections par cohorte, en les transposant dans un modèle économétrique vecteur autorégressif (VAR). De cette façon, la méthode des composantes se dote d'un cadre stochastique qui étend son envergure. Le potentiel de cette perspective est illustré à travers l'exemple d'une projection d'effectifs de population scolaire. Il met en valeur une série d'équations qui permet de vérifier la validité de plusieurs choix de modélisations habituellement utilisées dans le domaine de la prévision.

**English abstract, p. 170**

### **PROLÉGOMÈNES**

#### **Une perspective économétrique**

**L'**objet de cet article est de démontrer qu'il est possible de recourir à l'analyse des séries chronologiques afin de sous-tendre un nouveau type de projections démographiques incluant des intervalles de confiance, aussi grands soient-ils. Bien que les modèles autorégressifs servant à effectuer des prévisions soient bien connus des démographes, cette application particulière n'a encore jamais été tentée. Après avoir présenté notre cadre d'analyse et détaillé la structure de nos données dans une perspective économétrique puis démographique, nous passerons en seconde partie à l'application illustrative de l'analyse des séries chronologiques à la projection des effectifs de population scolaire d'un comté de l'État de Californie. Pour ce faire, l'utilisation de la méthode des composantes, par le biais des ratios de survie scolaire, mettra en évidence l'apport des séries chronolo-

---

\* Population Studies Center, Université de Pennsylvanie, Philadelphie, États-Unis.

giques à la projection démographique, de même que dans son extension stochastique. En conclusion, nous proposerons, à partir de ces résultats, des pistes de recherche prometteuses dans le domaine de la prévision démographique.

Pour commencer, nous présentons dans la *figure 1* une structure de données stylisées dans laquelle les rangées sont marquées par les indices  $\{t - T, \dots, t - 2, t - 1, t, t + 1, t + 2, \dots\}$  afin d'indiquer la nature temporelle de cet échafaudage. On observe les données pendant  $T + 1$  périodes dont la plus récente est  $t$ . Les lignes dénotées par  $t + \dots$  signifient l'avenir. Ainsi, on ignore les valeurs des cellules (grisâtres) qui s'y rapportent. Il s'agit de déceler des relations parmi les observations faites antérieurement dans le but de proposer un modèle qui pourra servir à prévenir l'évolution de ces valeurs.

**FIGURE 1** Quelques colonnes de données rangées par période

	1	2	3	...	$j$	...	$J$
$t - T$	$y_{t-T,1}$	$y_{t-T,2}$	$y_{t-T,3}$	...	$y_{t-T,j}$	...	$y_{t-T,J}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	...	$\vdots$	...	$\vdots$
$t - 2$	$y_{t-2,1}$	$y_{t-2,2}$	$y_{t-2,3}$	...	$y_{t-2,j}$	...	$y_{t-2,J}$
$t - 1$	$y_{t-1,1}$	$y_{t-1,2}$	$y_{t-1,3}$	...	$y_{t-1,j}$	...	$y_{t-1,J}$
$t$	$y_{t,1}$	$y_{t,2}$	$y_{t,3}$	...	$y_{t,j}$	...	$y_{t,J}$
$t + 1$	$y_{t+1,1}$	$y_{t+1,2}$	$y_{t+1,3}$	...	$y_{t+1,j}$	...	$y_{t+1,J}$
$t + 2$	$y_{t+2,1}$	$y_{t+2,2}$	$y_{t+2,3}$	...	$y_{t+2,j}$	...	$y_{t+2,J}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	...	$\vdots$	...	$\vdots$

Source : figure et notations de l'auteur.

Du point de vue de l'économétricien, il y a une grande classe de modèles qui se prêtent à l'analyse de ces séries chronologiques, dont la prévision. Sans restreindre la généralité, nous nous concentrerons sur le système d'équations suivant :

$$y_t = \alpha_0 + A_1 y_{t-1} + \dots + A_\ell y_{t-\ell} + \dots + A_p y_{t-p} + u_t$$

(1)

où pour  $1 \leq \ell \leq p \ll T$ , les  $\mathbf{y}$  sont  $p + 1$  vecteurs de longueur  $J$ , le nombre de colonnes dans la figure 1 :

$$\mathbf{y}_t = \begin{bmatrix} y_{1,t} \\ y_{2,t} \\ y_{3,t} \\ \vdots \\ y_{J,t} \end{bmatrix} \quad \mathbf{y}_{t-1} = \begin{bmatrix} y_{1,t-1} \\ y_{2,t-1} \\ y_{3,t-1} \\ \vdots \\ y_{J,t-1} \end{bmatrix} \quad \mathbf{y}_{t-\ell} = \begin{bmatrix} y_{1,t-\ell} \\ y_{2,t-\ell} \\ y_{3,t-\ell} \\ \vdots \\ y_{J,t-\ell} \end{bmatrix} \quad \mathbf{y}_{t-p} = \begin{bmatrix} y_{1,t-p} \\ y_{2,t-p} \\ y_{3,t-p} \\ \vdots \\ y_{J,t-p} \end{bmatrix} \quad (2)$$

$\alpha_0$  est un vecteur de constantes, également de longueur  $J$  :

$$\alpha_0 = \begin{bmatrix} \phi_1^0 \\ \phi_2^0 \\ \phi_3^0 \\ \vdots \\ \phi_J^0 \end{bmatrix} \quad (3)$$

et nous avons  $p$  matrices carrées de dimension  $J \times J$  qui relient les colonnes — en principe, chaque série à toutes les autres — jusqu'au décalage  $p$  :

$$A_1 = \begin{bmatrix} \phi_{1,1}^1 & \phi_{1,2}^1 & \phi_{1,3}^1 & \cdots & \phi_{1,J}^1 \\ \phi_{2,1}^1 & \phi_{2,2}^1 & \phi_{2,3}^1 & \cdots & \phi_{2,J}^1 \\ \phi_{3,1}^1 & \phi_{3,2}^1 & \phi_{3,3}^1 & \cdots & \phi_{3,J}^1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \phi_{J,1}^1 & \phi_{J,2}^1 & \phi_{J,3}^1 & \cdots & \phi_{J,J}^1 \end{bmatrix} \quad A_\ell = \begin{bmatrix} \phi_{1,1}^\ell & \phi_{1,2}^\ell & \phi_{1,3}^\ell & \cdots & \phi_{1,J}^\ell \\ \phi_{2,1}^\ell & \phi_{2,2}^\ell & \phi_{2,3}^\ell & \cdots & \phi_{2,J}^\ell \\ \phi_{3,1}^\ell & \phi_{3,2}^\ell & \phi_{3,3}^\ell & \cdots & \phi_{3,J}^\ell \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \phi_{J,1}^\ell & \phi_{J,2}^\ell & \phi_{J,3}^\ell & \cdots & \phi_{J,J}^\ell \end{bmatrix} \quad A_p = \begin{bmatrix} \phi_{1,1}^p & \phi_{1,2}^p & \phi_{1,3}^p & \cdots & \phi_{1,J}^p \\ \phi_{2,1}^p & \phi_{2,2}^p & \phi_{2,3}^p & \cdots & \phi_{2,J}^p \\ \phi_{3,1}^p & \phi_{3,2}^p & \phi_{3,3}^p & \cdots & \phi_{3,J}^p \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \phi_{J,1}^p & \phi_{J,2}^p & \phi_{J,3}^p & \cdots & \phi_{J,J}^p \end{bmatrix} \quad (4).$$

Quant aux erreurs, elles sont distribuées en tant qu' $\mathbf{u}_t \sim N(0, \Sigma_t)$ , avec la matrice variance-covariance  $\Sigma_t$ , également de dimension  $J \times J$ , typiquement de la forme

$$\Sigma_t = \begin{bmatrix} \sigma_{u_1}^2 & 0 & 0 & \cdots & 0 \\ 0 & \sigma_{u_2}^2 & 0 & \cdots & 0 \\ 0 & 0 & \sigma_{u_3}^2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \sigma_{u_J}^2 \end{bmatrix}. \quad (5)$$

Par exemple, si on substitue les contraintes

$$A_1^{AR} = \begin{bmatrix} \phi_{1,1}^1 & 0 & 0 & \cdots & 0 \\ 0 & \phi_{2,2}^1 & 0 & \cdots & 0 \\ 0 & 0 & \phi_{3,3}^1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \phi_{J,J}^1 \end{bmatrix} \quad A_\ell^{AR} = \begin{bmatrix} \phi_{1,1}^\ell & 0 & 0 & \cdots & 0 \\ 0 & \phi_{2,2}^\ell & 0 & \cdots & 0 \\ 0 & 0 & \phi_{3,3}^\ell & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \phi_{J,J}^\ell \end{bmatrix} \quad A_p^{AR} = \begin{bmatrix} \phi_{1,1}^p & 0 & 0 & \cdots & 0 \\ 0 & \phi_{2,2}^p & 0 & \cdots & 0 \\ 0 & 0 & \phi_{3,3}^p & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \phi_{J,J}^p \end{bmatrix} \quad (6)$$

à l'équation (4), on réalise  $J$  équations pour les processus  $AR(p)$ , auquel cas la prévision des valeurs futures au sein de chaque colonne ne s'appuie que sur les données antérieures à de la même colonne. En revanche, si on veut se servir de toutes les données antérieures, toutes variables confondues — c'est-à-dire que le passé de chaque variable peut influencer la prévision de toutes les autres — on se retrouve avec les équations (1)-(5) et, en particulier, avec l'équation (4) complètement non contrainte, au moins dans la mesure où  $T \gg J$ , afin que tous les paramètres soient identifiés<sup>1</sup>. C'est un VAR ou ARV, un modèle vecteur autorégressif (Enders, 2004 : 264-272), encore appelé autorégression vectorielle (Gujarati, 2004 : 838-845).

Entre les deux extrêmes, les séries isolées de l'équation (6) et le « tout empiète sur tout » de l'équation (4), il existe une gamme de modèles qui ont été développés dans le cadre des modèles SVAR ou VARS, vecteur autorégressif structurel (Sims, 1986 ; e. g. Desnoyers, 2001). Bon nombre de théories et de débats ont marqué l'évolution de ces modèles et ont donné lieu à maintes discussions sur les circonstances dans lesquelles il faut, ou il ne faut pas, imposer des contraintes à la matrice de l'équation (4) (Stock et Watson, 2001 ; Favero, 2007). Nous n'entrerons pas dans ces débats, car ils s'encombrent d'un discours économique qui ne nous concerne pas, sauf pour constater ici qu'il s'agit de distinguer les variables qui sont exogènes par rapport aux autres. Dans l'équation (4), toutes les variables sont mutuellement endogènes.

La notation sera adaptée plus bas de façon à écarter les variables exogènes des variables endogènes, à l'instar d'un modèle VARX (Bierens, 2004), avec les vecteurs  $\mathbf{x}$  pour les variables exogènes et les vecteurs  $\mathbf{y}$  réservés aux variables endogènes. Pour l'instant, il suffit de considérer que les vecteurs  $\mathbf{y}_1$  et  $\mathbf{y}_2$  restent endogènes, en ce sens que chacun empiète sur l'autre, et que tous les deux dépendent des autres vecteurs,  $\mathbf{y}_3, \dots, \mathbf{y}_p, \dots, \mathbf{y}_p$  mais qu'ils ne causent pas, au sens de Granger (1969), cette deuxième série de variables. Par conséquent, on précise les matrices à la place de celles de

1. On a constaté dès le début de ce modèle que le nombre des paramètres s'accroissait avec le carré du nombre des variables (colonnes de la *figure 1*) pour que les degrés de liberté soient rapidement épuisés (Sims, 1980 : 16). Il y a  $p \times J^2$  paramètres dans l'équation (4) (et  $2 \times J$ , de surcroît, dans les équations [3] et [5]) contre  $T \times J$  points de données observées dans la *figure 1*. Par conséquent, il faut que  $(T - p) > (p \times J) + 2$  pour pouvoir estimer tous ces paramètres. Ce problème d'identification a eu tendance à se résoudre par rapport à la précision de la longueur des décalages  $p$ , et devait se distinguer par la suite des problèmes d'identification abordés dans les manuels (e. g. Enders, 2004 : 269-272, 299-301) des soi-disant paramètres structurels.

$$A_1^S = \begin{bmatrix} \phi_{1,1}^1 & \phi_{1,2}^1 & \phi_{1,3}^1 & \cdots & \phi_{1,J}^1 \\ \phi_{2,1}^1 & \phi_{2,2}^1 & \phi_{2,3}^1 & \cdots & \phi_{2,J}^1 \\ 0 & 0 & \phi_{3,3}^1 & \cdots & \phi_{3,J}^1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \phi_{J,3}^1 & \cdots & \phi_{J,J}^1 \end{bmatrix} \quad A_2^S = \begin{bmatrix} \phi_{1,1}^\ell & \phi_{1,2}^\ell & \phi_{1,3}^\ell & \cdots & \phi_{1,J}^\ell \\ \phi_{2,1}^\ell & \phi_{2,2}^\ell & \phi_{2,3}^\ell & \cdots & \phi_{2,J}^\ell \\ 0 & 0 & \phi_{3,3}^\ell & \cdots & \phi_{3,J}^\ell \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \phi_{J,3}^\ell & \cdots & \phi_{J,J}^\ell \end{bmatrix} \quad A_p^S = \begin{bmatrix} \phi_{1,1}^p & \phi_{1,2}^p & \phi_{1,3}^p & \cdots & \phi_{1,J}^p \\ \phi_{2,1}^p & \phi_{2,2}^p & \phi_{2,3}^p & \cdots & \phi_{2,J}^p \\ 0 & 0 & \phi_{3,3}^p & \cdots & \phi_{3,J}^p \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \phi_{J,3}^p & \cdots & \phi_{J,J}^p \end{bmatrix} \quad (7)$$

l'équation (4). Ce schéma de base — à remarquer que le coin gauche en bas de chaque matrice ne se compose que des éléments nuls — est typique des dessins qui viennent structurer les VARS, autrement débridés.

Ce qui nous intéresse ici, c'est que la caractérisation exogène d'un tel ensemble de variables et les contraintes conséquentes à la façon de l'équation (7) sont censées découler de la théorie économique (*cf.* Sims, 1980). Or, celles-ci n'ont rien à voir avec l'ordre des colonnes dans la *figure 1*. Les indices  $J = 1, \dots, J$  ne servent qu'à indiquer les variables au sens nominal et sont sans signification ordinale. On peut les permuter pourvu qu'on se souvienne à quel indice correspond chaque variable. Par contre, on ne peut pas permuter les lignes du tableau de la *figure 1* sans perdre le sens temporel des données.

## Une perspective démographique

Dans l'optique d'un démographe, une structure de données comme celle de la *figure 1* peut représenter une deuxième dimension de temps, soit de l'âge ou, plus généralement, de la durée. Dans cette perspective, les diagonales d'un tel tableau prennent une signification, comme le montre la *figure 2*, où l'on suppose que la distance entre chaque période (chaque rangée) est identique à celle de chaque âge ou durée (chaque colonne). L'exposant,  $k = (t - T - J + 1), \dots, t - T, \dots, t - 2, t - 1, t, t + 1, t + 2, \dots$ , marque, pour chaque diagonale, le temps d'entrée dans la première colonne de la structure. Ce sont les *cohortes* qui vieillissent du haut, à gauche de la figure, vers le bas, à droite, avec le déroulement du temps. Comme la vie est relativement longue, de nombreuses régularités par âge apparaissent au travers de son cours, c'est-à-dire, au sein d'une cohorte. Ainsi, la projection par cohorte est un outil très puissant pour la prévision chez les démographes<sup>2</sup>. Ceci est explicite dans les modèles canoniques de la

2. « Les démographes distinguent habituellement les prévisions, des projections ou perspectives. En regardant de plus près, on y trouve, plutôt, une précaution de langage qu'une différence de nature... Prévision et projection s'appuyent [*sic*] toutes deux sur le passé pour pronostiquer l'avenir » (Le Bras, 1987 : 13-14). Dans la suite de l'article, je ferai dorénavant référence à Le Bras en utilisant les deux termes de manière interchangeable.

projection d'une population par âge au moyen des matrices de Leslie (Namboodiri, 1991 : 160-172, 184) et d'autres (Ledent, 1995). C'est la colonne vertébrale du modèle fondamental des prévisions démographiques qui s'appuient sur la méthode des composantes (Leridon, 2003 : 630), ce que Le Bras (2005 : 174) a appelé « le noyau dur de la discipline ». Par exemple, on peut prévoir l'avenir du tabagisme (Wang, 2008 : 73-110) parce que le niveau de tabagisme d'une cohorte se fige par rapport au seuil d'initiation chez les adolescents et les jeunes adultes (Inserm, 2003 : 11), et parce que la mortalité s'ensuit forcément, on peut l'anticiper également (Wang et Preston, 2009). C'est probablement la raison pour laquelle seuls les démographes se permettent de faire des prévisions à longue échéance, dans un horizon de dix ans ou plus encore (Granger, 2007 : 6).

**FIGURE 2** Quelques données rangées par durée, période et (sur les diagonales) cohorte

		Durée						
		1	2	3	...	$j$	...	$J$
Période	$t - T$	$w_{t-T,1}^{t-T}$	$w_{t-T,2}^{t-T-1}$	$w_{t-T,3}^{t-T-2}$	...	$w_{t-T,j}^{t-T-j+1}$	...	$w_{t-T,J}^{t-T-J+1}$
	$\vdots$	$\vdots$	$\vdots$	$\vdots$	...	$\vdots$	...	$\vdots$
	$t - 2$	$w_{t-2,1}^{t-2}$	$w_{t-2,2}^{t-3}$	$w_{t-2,3}^{t-4}$	...	$w_{t-2,j}^{t-j-1}$	...	$w_{t-2,J}^{t-J-1}$
	$t - 1$	$w_{t-1,1}^{t-1}$	$w_{t-1,2}^{t-2}$	$w_{t-1,3}^{t-3}$	...	$w_{t-1,j}^{t-j}$	...	$w_{t-1,J}^{t-J}$
	$t$	$w_{t,1}^t$	$w_{t,2}^{t-1}$	$w_{t,3}^{t-2}$	...	$w_{t,j}^{t-j+1}$	...	$w_{t,J}^{t-J+1}$
	$t + 1$	$w_{t+1,1}^{t+1}$	$w_{t+1,2}^t$	$w_{t+1,3}^{t-1}$	...	$w_{t+1,j}^{t-j+2}$	...	$w_{t+1,J}^{t-J+2}$
	$t + 2$	$w_{t+2,1}^{t+2}$	$w_{t+2,2}^{t+1}$	$w_{t+2,3}^t$	...	$w_{t+2,j}^{t-j+3}$	...	$w_{t+2,J}^{t-J+3}$
	$\vdots$	$\vdots$	$\vdots$	$\vdots$	...	$\vdots$	...	$\vdots$

Source : figure et notations de l'auteur.

Alho et Spencer (2005 : 201) ont remarqué que les statisticiens probabilistes ont développé les modèles linéaires stationnaires au cours des années 1920, 1930 et 1940, en même temps que les démographes construisaient des projections originales à l'aide de la méthode des cohortes. Les auteurs ont abordé ces deux sujets en profondeur (*inter alia*, p. 180-186, 198-238). Ils ont estimé plusieurs modèles mixtes intégrés autorégressifs et de moyennes mobiles (ARIMA) pour les données démographiques clas-

sées d'après la *figure 2* — les taux de fécondité et de mortalité, rangés par année et par âge — et ont déterminé qu'il n'y avait pas de séries chronologiques stationnaires, quel que soit l'âge (p. 207). En outre, et davantage intéressant, les auteurs ont constaté que le modèle VAR représentait une généralisation de plusieurs modèles AR, à l'instar de l'équation (4) par opposition à l'équation (6) (p. 204) et que la même algèbre matricielle de la projection par cohorte s'employait d'une façon différente par rapport à l'analyse des séries chronologiques (p. 181). Cependant, les deux littératures n'ont pas toujours évolué singulièrement; elles ont eu à s'entrecroiser plusieurs fois. On recense ainsi plusieurs illustrations de ce fait (*inter alia*, Butz et Ward, 1979; McDonald, 1980; Lee et Carter, 1992; Wachter, 1992 : 81; Lee et Tuljapurkar, 1994; Renshaw et Haberman, 2003) et on pourrait multiplier les exemples (*cf.* Booth, 2006).

De cette manière, « l'astuce » élaborée ci-dessus, le changement du sens des colonnes et le remaniement dans la foulée de la perspective sur la prévision, n'apparaît pas en soi comme quelque chose d'important. Pourtant, il s'avère que le modèle VAR constitue un cadre d'analyse très fructueux pour la compréhension, voire l'amélioration des dynamiques des modèles de projection démographique d'un usage courant. De la même façon, la structure par cohorte peut donner, par le biais des modèles VAR, une « mémoire » inédite aux séries temporelles. Nous en démontrons le bien-fondé par rapport au problème de la prévision de l'inscription scolaire.

### EXEMPLE D'UNE SYNTHÈSE : LA PROJECTION DE L'EFFECTIF SCOLAIRE

Que ce soit au niveau national (UNESCO, 2009), provincial (Gouvernement du Québec, 2009) ou local (Morrison, 2000; Projet..., 2007) — même pour une école particulière (Lamy, 2004) — la projection (voire la prévision) de l'effectif scolaire s'appuie la plupart du temps sur une version de la méthode de la « survie de la cohorte ». L'idée est assez simple :

Posons que les éléments  $\{w_{t,j}^k\}$  de la *figure 2* expriment les effectifs d'élèves par classe afin que dans l'année  $t$ , il y ait  $w_{t,j}^{t-j+1}$  étudiants au niveau  $j$ . Pour  $i \leq t$ , ces effectifs sont constatés mais pour  $i > t$ , il s'agit de l'avenir, et donc, de la prévision. Conséquemment, l'effectif dans la classe dans l'année  $t+1$  aurait dû être prédit en fonction de l'effectif dans la classe précédente ( $j-1$ ) dans l'année antérieure ( $t$ ). On écrit



$$w_{t+1,j}^{t-j+2} = {}_{j-1}^j \varphi[\rho] \times w_{t,j-1}^{t-j+2} \quad (8)$$

où les  ${}_{j-1}^j \varphi[\rho]$  représentent les « rapports de survie scolaire » ou les « rapports de progression » (Jacoby, 1959 : 28), par analogie avec le terme courant en anglais « *grade-progression ratios* » (e. g. Siegel, 2002 : 516). Ils se fondent sur les relations qui ont existé antérieurement. Ainsi, pour une seule période passée,  $\rho = 1$ ,

$${}_{j-1}^j \varphi[1] = \frac{w_{t,j}^{t-j+1}}{w_{t-1,j-1}^{t-j+1}}.$$

En combinaison avec l'équation (8), on se sert du rapport qui a fait le glissement d'un niveau ( $j - 1$ ) à un autre ( $j$ ) dans la cohorte précédente ( $t - j + 1$ ) pour prévoir ce qui va arriver au sein de la cohorte  $t - j + 2$  entre les mêmes niveaux.

Dans le cas où l'on veut se servir de deux périodes passées ( $\rho = 2$ ) pour estimer les rapports de survie, on constate la moyenne géométrique des deux rapports antérieurs :

$${}_{j-1}^j \varphi[2] = \left( \frac{w_{t,j}^{t-j+1}}{w_{t-1,j-1}^{t-j+1}} \times \frac{w_{t-1,j}^{t-j}}{w_{t-2,j-1}^{t-j}} \right)^{\frac{1}{2}}$$

et, en général, on peut écrire

$${}_{j-1}^j \varphi[\rho] = \left( \prod_{\ell=1}^{\rho} \frac{w_{t-\ell+1,j}^{t-\ell-j+2}}{w_{t-\ell,j-1}^{t-\ell-j+2}} \right)^{\frac{1}{\rho}} \quad (9)$$

pour la moyenne géométrique simple de  $\rho$  rapports passés.

Pour relier les éléments de la *figure 1* à ceux de la *figure 2*, on définit :

$$y_{t,j} \equiv \ln w_{t,j}^{t-j+1} \quad (10)$$

de telle manière que le système d'équations (8)-(9) soit rendu linéaire. Pour  $\rho = 1$ , on prévoit :

$$\begin{aligned} w_{t+1,j}^{t-j+2} &= \frac{w_{t,j}^{t-j+1}}{w_{t-1,j-1}^{t-j+1}} \times w_{t,j-1}^{t-j+2} \\ \ln w_{t+1,j}^{t-j+2} &= \ln \frac{w_{t,j}^{t-j+1}}{w_{t-1,j-1}^{t-j+1}} + \ln w_{t,j-1}^{t-j+2} \\ \ln w_{t+1,j}^{t-j+2} &= \ln w_{t,j}^{t-j+1} - \ln w_{t-1,j-1}^{t-j+1} + \ln w_{t,j-1}^{t-j+2} \end{aligned}$$

et, en conséquence de l'identité (10) :

$$y_{t+1,j} = y_{t,j} - y_{t-1,j-1} + y_{t,j-1}.$$

Pour  $\rho = 2$ , il s'agit de

$$\begin{aligned} w_{t+1,j}^{t-j+2} &= \left( \frac{w_{t,j}^{t-j+1}}{w_{t-1,j-1}^{t-j+1}} \times \frac{w_{t-1,j}^{t-j}}{w_{t-2,j-1}^{t-j}} \right)^{\frac{1}{2}} \times w_{t,j-1}^{t-j+2} \\ \ln w_{t+1,j}^{t-j+2} &= \frac{1}{2} \ln \left( \frac{w_{t,j}^{t-j+1}}{w_{t-1,j-1}^{t-j+1}} \times \frac{w_{t-1,j}^{t-j}}{w_{t-2,j-1}^{t-j}} \right) + \ln w_{t,j-1}^{t-j+2} \\ w_{t+1,j}^{t-j+2} &= \frac{1}{2} \ln w_{t,j}^{t-j+1} - \frac{1}{2} \ln w_{t-1,j-1}^{t-j+1} + \frac{1}{2} \ln w_{t-1,j}^{t-j} - \frac{1}{2} \ln w_{t-2,j-1}^{t-j} + \ln w_{t,j-1}^{t-j+2} \\ y_{t+1,j} &= \frac{1}{2} y_{t,j} - \frac{1}{2} y_{t-1,j-1} + \frac{1}{2} y_{t-1,j} - \frac{1}{2} y_{t-2,j-1} + y_{t,j-1}. \end{aligned}$$

De cette manière, on aboutit à la forme générale  $\rho$ ,

$$y_{t+1,j} = \frac{1}{\rho} \sum_{\ell=1}^{\rho} y_{t-\ell+1,j} - \frac{1}{\rho} \sum_{\ell=1}^{\rho} y_{t-\ell,j} + y_{t,j-1}. \quad (11)$$

Soulignons que l'équation (11) représente effectivement un ensemble d'équations dans la mesure où, pour  $J$  niveaux (classes) dans un système scolaire, cette projection se répète à  $J - 1$  reprises. De plus, étant entendu que l'effectif pronostiqué au niveau  $j$  pour l'année  $t + 1$  ( $y_{t+1,j}$ ) s'implique subséquentment dans la projection de l'effectif au niveau  $j + 1$  pour l'année  $t + 2$  ( $y_{t+2,j+1}$ ), et ainsi de suite, nous rappellerons qu'une telle combinaison linéaire pour la projection « se prête alors facilement au calcul matriciel » (Le Bras, 2005 : 188). En particulier, on précise pour l'équation (11) et tous les  $j$ ,

$$y_{t+1} = \alpha_0 + A_1^{M;\rho} y_t + \dots + A_{\ell}^{M;\rho} y_{t-\ell+1} + \dots + A_p^{M;\rho} y_{t-\rho} \quad (12)$$

où les  $y$  sont  $p + 1 (= \rho + 2)$  vecteurs de longueur  $J$ ,

$$y_{t+1} = \begin{bmatrix} y_{1,t+1} \\ y_{2,t+1} \\ y_{3,t+1} \\ \vdots \\ y_{J,t+1} \end{bmatrix} \quad y_t = \begin{bmatrix} y_{1,t} \\ y_{2,t} \\ y_{3,t} \\ \vdots \\ y_{J,t} \end{bmatrix} \quad y_{t-\ell+1} = \begin{bmatrix} y_{1,t-\ell+1} \\ y_{2,t-\ell+1} \\ y_{3,t-\ell+1} \\ \vdots \\ y_{J,t-\ell+1} \end{bmatrix} \quad y_{t-p} = \begin{bmatrix} y_{1,t-p} \\ y_{2,t-p} \\ y_{3,t-p} \\ \vdots \\ y_{J,t-p} \end{bmatrix} \quad (13)$$

et où il y a  $p = \rho + 1$  matrices de type  $A^{M;p}$  (moyenne simple de  $\rho$  rapports de progression) :

$$A_1^{M;p} = \begin{bmatrix} 0 & 0 & 0 & \dots & 0 \\ 1 & \rho^{-1} & 0 & \dots & 0 \\ 0 & 1 & \rho^{-1} & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \rho^{-1} \end{bmatrix} \quad A_\ell^{M;p} = \begin{bmatrix} 0 & 0 & 0 & \dots & 0 \\ -\rho^{-1} & \rho^{-1} & 0 & \dots & 0 \\ 0 & -\rho^{-1} & \rho^{-1} & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \rho^{-1} \end{bmatrix} \quad A_p^{M;p} = \begin{bmatrix} 0 & 0 & 0 & \dots & 0 \\ -\rho^{-1} & 0 & 0 & \dots & 0 \\ 0 & -\rho^{-1} & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 0 \end{bmatrix} \quad (14)$$

pour  $\ell = 1, \dots, \rho^3$  (ainsi, pour  $\rho = 2$ , il n'y a que deux matrices). Il convient de remarquer que bien qu'en passant des équations (8) et (9) à l'équation (11) on se débarrasse des exposants qui représentent les cohortes, les matrices de l'équation (14) retiennent fidèlement le système de projection par la méthode des composantes, c'est-à-dire par les rapports de survie scolaire appliqués aux cohortes. Notons également que ces équations s'accordent parfaitement avec l'équation (9), il faut stipuler ici que

$$\alpha_0 = [0]_{J \times 1} \quad (15)$$

mais on retient ce vecteur pour les ordonnées à l'origine parce qu'on s'en sert plus tard.

Le système d'équations (12)-(14) ressemble beaucoup à celui des équations (1)-(5), sauf sur les quelques points qui suivent :

- les erreurs ne figurent plus;
- le vecteur à gauche du signe d'égalité dans l'équation (12) a rapport à la période  $t + 1$ , c'est-à-dire à l'avenir (et ainsi tous les décalages à droite se déplacent d'une période par rapport à ceux dans l'équation [1]);
- et enfin dans l'équation (14), il n'y a que des paramètres fixes, tandis que même avec beaucoup de contraintes sur les paramètres dans l'équation (3), à l'instar de celles de l'équation (6), il reste des paramètres libres.

Ces différences s'expliquent par une différence d'optique. La nouvelle perspective temporelle de l'équation (12) face à celle de l'équation (1) n'est pas une défaillance; elle découle directement de la définition canonique d'un rapport de survie scolaire et de son application aux données disponibles. Dans cette optique, il ne s'agit que de préciser  $\rho$ , le nombre de périodes passées qui figurent dans le calcul d'un rapport de survie scolaire et d'avancer les effectifs observés par classe dans l'année  $t$ , au moyen de

3. La prévision est pour les classes  $1 < j \leq J$ . Tous les éléments dans le coin supérieur gauche des matrices de l'équation (14) ont été fixés à nul parce que la valeur prévue de  $y_{1,p}$ , le *kindergarten*, qui est également la première classe au sein du système scolaire, va dépendre des séries de naissances, encore inconnues.

l'équation (8). Ayant réécrit cette équation afin qu'elle ressemble à un modèle vecteur autorégressif, on constate que ce processus s'appuie effectivement sur un passé un peu plus profond, relié au futur par les paramètres de l'équation (14). Pourtant, ces paramètres sont fixes, conditionnels de  $\rho$  (et supposent que la moyenne dans l'équation [9] soit simple); ainsi, le système d'équations (12)-(14) est un système de projection et non un système d'estimation.

Dans la perspective économétrique, la prévision se déroule en deux étapes. On précise un modèle (voire plusieurs modèles) fondé sur les données dans l'équation (3) ainsi que les contraintes sur la répartition des erreurs, telles que celles de l'équation (5), afin d'estimer les paramètres dans l'équation (4) (ou les équations de contrainte, telles que [6] ou [7]). Après avoir établi ces paramètres, on peut commencer à faire des prévisions, d'abord pour l'année  $t + 1$ , en mettant à jour tous les décalages pour que l'équation (1) se transforme en équation de prévision comme l'équation (12).

### **Des intervalles de confiance pour les projections des effectifs scolaires**

On fait dériver les variances des erreurs de prévision des séries chronologiques sous l'hypothèse simplificatrice que les coefficients matriciels des équations (3) et (4) sont connus (*e. g.* Enders, 2004 : 279-280, 291)<sup>4</sup>. Avec la projection par la méthode des composantes, on peut en effet renverser le processus : on pose les coefficients par le biais des équations (14) et (15); on déplace en temps l'équation (12); et on retrouve des erreurs du passé sous l'hypothèse que le modèle pour faire les projections dans le futur a duré longtemps :

$$\mathbf{u}_t \equiv \mathbf{y}_t - (\alpha_0 + A_1^{M;\rho} \mathbf{y}_{t-1} + \dots + A_\ell^{M;\rho} \mathbf{y}_{t-\ell} + \dots + A_p^{M;\rho} \mathbf{y}_{t-p}). \quad (16)$$

Par rapport à celles de l'équation (1), ces erreurs sont complètement empiriques dans le sens où il n'existe pas de supposition sur leur répartition. Ainsi, on ne connaît pas réellement les propriétés statistiques des intervalles de confiance fondés sur ces erreurs de prévision. Malgré tout, les intervalles de confiance dans les projections démographiques non

4. Autrement, il faut considérer simultanément le manque de précision qui est inhérent aux estimateurs de ces coefficients-ci; c'est la raison pour laquelle les manuels préconisent les séries assez longues pour qu'on puisse croire au modèle sélectionné et construire aux intervalles de confiance (Enders, 2004 : 82).

stochastiques restent une idée qui séduit (*e. g.* Tuljapurkar, 1992; Duchêne, 1999) et, jusqu'ici, l'évaluation des erreurs de prévision dans la projection des effectifs scolaires n'existe que par celles observées après l'écoulement des années (*e. g.* Spar, 1994; Grip, 2004, 2009).

Pour saisir dans le cadre des modèles VAR très contraints les apports et les limites de la projection par la méthode des composantes, nous considérerons ici deux modèles de prévision très élémentaires qui s'appliquent aux données d'un comté, tirées d'une façon peu méthodique de State of California (2008) pour les années 1974 à 2008 et portant sur 15 000 élèves par classe. On fixe le nombre des classes à  $J = 9$ , soit un *kindergarten* ( $j = 1$ ) suivi par les classes 1 à 8. Pour le premier modèle, on pose  $\rho = 1$ , c'est-à-dire qu'on ne se sert que du rapport de progression (pour chaque classe) le plus récent en faisant les projections. Pour le second modèle, il s'agit de  $\rho = 2$ , on fait une moyenne géométrique des rapports de progression de 2006 à 2007 et de 2007 à 2008. En particulier, pour  $\rho = 1$  on précise

$$A_1^{M;1} = \begin{bmatrix} \phi_{1,1}^1 & 0 & 0 & \cdots & 0 \\ 1 & 1 & 0 & \cdots & 0 \\ 0 & 1 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix}_{9 \times 9} \quad A_2^{M;1} = \begin{bmatrix} \phi_{1,1}^2 & 0 & 0 & \cdots & 0 \\ -1 & 0 & 0 & \cdots & 0 \\ 0 & -1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix}_{9 \times 9} \quad (17)$$

et pour  $\rho = 2$ ,

$$A_1^{M;2} = \begin{bmatrix} \phi_{1,1}^1 & 0 & 0 & \cdots & 0 \\ 1 & \frac{1}{2} & 0 & \cdots & 0 \\ 0 & 1 & \frac{1}{2} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \frac{1}{2} \end{bmatrix}_{9 \times 9} \quad A_2^{M;2} = \begin{bmatrix} \phi_{1,1}^2 & 0 & 0 & \cdots & 0 \\ -\frac{1}{2} & \frac{1}{2} & 0 & \cdots & 0 \\ 0 & -\frac{1}{2} & \frac{1}{2} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \frac{1}{2} \end{bmatrix}_{9 \times 9} \quad A_3^{M;2} = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 \\ -\frac{1}{2} & 0 & 0 & \cdots & 0 \\ 0 & -\frac{1}{2} & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix}_{9 \times 9} \quad (18).$$

On suppose que les effectifs pour  $j = 1$ , donc  $\{y_{1,t}\}$ , s'engendrent seulement grâce à un processus autorégressif d'ordre 2, ainsi

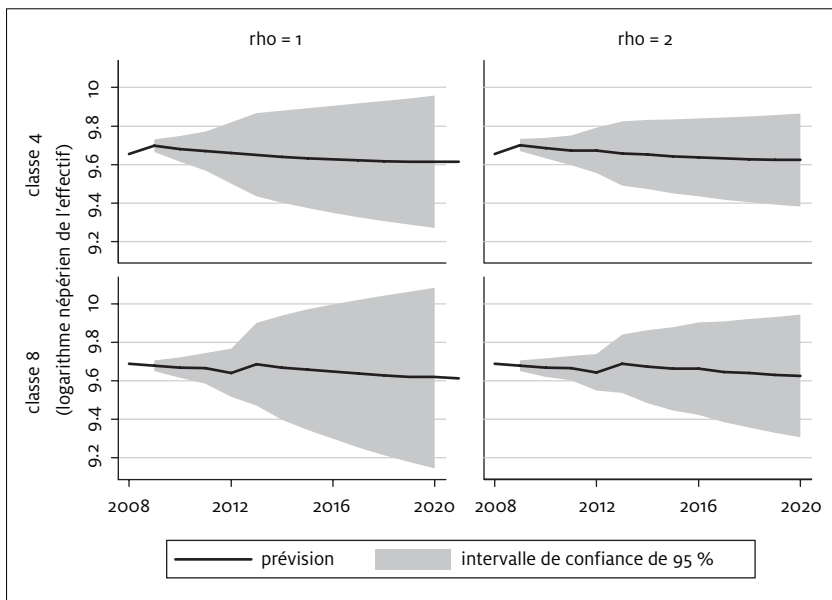
$$\alpha_0 = \begin{bmatrix} \phi_1^0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}_{9 \times 1} \quad (19)$$

également pour les deux modèles<sup>5</sup>.

5. Pour que les différences dans l'estimation des erreurs de prévision entre les deux modèles ne réfléchissent pas les différences secondaires, mais épineuses, dans l'estimation des séries pour  $\{y_{1,t}\}$  ainsi que pour  $\{u_{1,t}\}$ , on exige aussi que  $\phi_{1,t}^1 = \phi_{1,t}^2$  à travers les deux équations (17) et (18) et que l'estimation se réalise sur la même durée de données en dépit de la différence dans la longueur du décalage entre  $\rho = 1 [\Rightarrow \rho = 2]$  et  $\rho = 2 [\Rightarrow \rho = 3]$ .

Les résultats pour les classes 4 et 8 sont présentés dans la *figure 3*, avec  $\rho = 1$  à gauche et  $\rho = 2$  à droite. En raison de l'équation (10), l'ordonnée est en échelle logarithmique. Les équations (17) et (18) font ressortir l'atout de la méthode de projection des composantes par le biais des cohortes : ainsi, par exemple, l'augmentation de l'effectif dans la classe 4 qui est prévue en 2009 se reproduit dans les projections pour 2013 pour la classe 8, c'est-à-dire quatre ans plus tard, quand ces élèves ont avancé. Cet enchaînement n'existe pas dans le cas où les prévisions sont effectuées dans un système d'équations isolées, tel que celui précisé dans l'équation (6). Autrement dit, il ne peut se produire si les séries d'effectifs sont abordées classe par classe au sein d'un système dépourvu d'une structure de cohorte et déconnecté des modèles autorégressifs.

**FIGURE 3** Prévisions avec intervalle de confiance de 95 %, pour la classe 4 et la classe 8 de deux modèles (nombre de rapports de progression utilisés, soit  $\rho[\rho] = 1$  ou  $\rho[\rho] = 2$ )



Source : données de 1974 à 2008 pour un comté de Californie, State of California, 2008.

Quant aux estimations de l'erreur dans les prévisions : les racines carrées de la variance résiduelle estimées pour la classe 4 et la classe 8 s'avoisinent à  $\hat{\sigma}_{u4} \cong 0,016$  et  $\hat{\sigma}_{u8} \cong 0,014$ , c'est-à-dire une erreur moyenne de

1,5 %. Il s'avère qu'il n'y a aucune différence dans l'estimation de ces paramètres par rapport aux spécifications  $\rho = 1$  et  $\rho = 2^6$ . Par conséquent, les estimations des intervalles de confiance commencent, pour l'année 2009, avec ces valeurs (gonflées par le facteur 1,96 pour approcher le degré de confiance de 95 %) ... et se détériorent par la suite. La vitesse et l'envergure de cette détérioration sont fonction de la durée, de la classe et, fait intéressant, du choix de  $\rho$ , le nombre de rapports de progression historique qui figurent dans le calcul des rapports pour la projection.

Le fait que plus on se déplace vers l'avenir plus les erreurs de prévision augmentent est en accord avec le « principe d'incertitude croissante » (Léridon, 2003 : 630) et avec l'arithmétique : l'évolution des valeurs pour n'importe quelle série  $\{y_{t+1,j}, y_{t+2,j}, \dots, y_{t+n,j}\}$  est un processus dans lequel la prévision de chaque élément se fonde sur la valeur de la prévision précédente. Ainsi,  $\hat{y}_{t+2,j}$  s'appuie sur  $\hat{y}_{t+1,j}$  et  $\hat{y}_{t+n,j}$  sur  $\hat{y}_{t+n-1,j}$  de manière que les erreurs s'accumulent avec le temps (Enders, 2004 : 79-81).

Quant à la rapidité avec laquelle l'intervalle de confiance s'élargit : il est important de reconnaître que, bien que les équations du type de la (17) n'apparaissent que relier les séries contiguës, les valeurs antérieures — et les erreurs antérieures — se propagent en fait d'une série à l'autre au fil des années. Par exemple, avec  $\rho = 1$ , et la définition d'une différence,  $\Delta y_{t,j} = y_{t,j} - y_{t-1,j}$ , on peut réécrire le système d'équations (16), (13), (17) et (19) comme tel

$$\begin{aligned}\Delta y_{t,1} &= \phi_1^0 + (\phi_{1,1}^1 - 1)y_{t-1,1} + \phi_{1,1}^2 y_{t-2,1} + u_{t,1} \\ \Delta y_{t,2} &= \Delta y_{t-1,1} + u_{t,2} \\ \Delta y_{t,3} &= \Delta y_{t-1,2} + u_{t,3} \\ &\vdots\end{aligned}$$

donc  $\Delta y_{t-1,2} = \Delta y_{t-2,1} + u_{t-1,2}$  et  $\Delta y_{t-2,1} = \phi_1^0 + (\phi_{1,1}^1 - 1)y_{t-3,1} + \phi_{1,1}^2 y_{t-4,1} + u_{t-2,1}$  ; et, après substitution

$$\Delta y_{t,3} = \phi_1^0 + (\phi_{1,1}^1 - 1)y_{t-3,1} + \phi_{1,1}^2 y_{t-4,1} + u_{t-2,1} + u_{t-1,2} + u_{t,3}.$$

On pourrait répéter ce processus jusqu'à  $j = 5$  (la classe 4) et même  $j = J = 9$  (la classe 8) pour établir l'influence des erreurs à travers toutes les équations antécédentes. On remarque que les prévisions pour la classe 2

6. Les erreurs estimées sont pareilles pour les classes qui n'apparaissent pas dans la figure 3, sauf que la variance résiduelle pour le *kindergarten*,  $j = 1$ , qui ne dépend que de sa propre histoire, est plus élevée, soit environ de  $\hat{\sigma}_{\mu_1}^2 \approx 3,5 \times \hat{\sigma}_{\mu_{i>1}}^2$ .

(soit  $j = 3$ ), par exemple, deviennent une fonction croissante des erreurs, pas seulement celles de sa propre série ( $u_{t,3}$ ) et de la classe antérieure (c'est-à-dire la classe 1 ou  $j = 2$ ) qui figurent directement dans l'équation originale pour  $y_{t,3}$  (ou  $\Delta y_{t,3}$ ), mais également des paramètres et des erreurs de la série autorégressive pour le *kindergarten* (soit  $j = 1$ ). Tout se complique davantage quand on transcrit les valeurs comme  $y_{t-3,1}$  et  $y_{t-4,1}$  en fonction de leurs erreurs.

Par conséquent, les séries pour la classe 4 et pour la classe 8 puisent au début dans les effectifs antérieurs observés. À titre d'exemple, quand on fait le calcul pour la projection de  $y_{2009,9}$ , les valeurs pour  $y_{2008,8}$ ,  $y_{2008,9}$  et  $y_{2007,8}$  sont connues et non prévues ou projetées, et les erreurs sont de plus ou moins la même taille. Après quatre années, les prévisions pour la classe 4 ont épuisé l'information, dans le sens d'effectifs antérieurs observés, si fait qu'on ne tire plus que des erreurs du système. Par contre, les prévisions pour la classe 8 profitent des observations des effectifs réels pendant plusieurs années additionnelles.

D'ailleurs, on peut redéfinir  $\{y_{t,9}\}$  en fonction d'une plus grande quantité d'équations précédentes, ce qui signifie également une plus grande récolte d'erreurs (Enders, 2004 : 278-280). Pour cette raison, après la huitième année, quand toutes les informations ont été épuisées sauf pour les erreurs, l'effondrement est encore plus prononcé. On remarque pour les variances des erreurs de prévision une tendance à se stabiliser qui s'accorde avec les résultats des études de projections d'effectifs scolaires de Spar (1994; Demographics and Workforce, sans date), et peut-être aussi avec quelques résultats obtenus sur le plan théorique (Enders, 2004 : 80). Toutefois, ces derniers supposent que les séries sont stationnaires, or cette supposition ne s'accorde pas avec les coefficients précisés de 1, de  $-1$  et de  $-\frac{1}{2}$ , d'après les équations (17) et (18).

Enfin, on doit noter la grande différence existant entre les intervalles de confiance pour les spécifications  $\rho = 1$  et  $\rho = 2$ . Il semblerait que les erreurs de prévision s'atténuent avec la projection fondée sur la moyenne des deux rapports de progression les plus récents, comparé à l'emploi du plus récent rapport uniquement. Cependant, nous devons rappeler qu'il n'existe aucune différence entre les deux modèles par rapport à l'estimation de la variance résiduelle. Par conséquent, les différences dans les erreurs de prévision ne résultent que des différences entre les matrices des équations (17) et (18). Peut-on prouver de cette façon la supériorité des systèmes par lesquels on estime les rapports de projection d'une moyenne plus inclusive ? On ne peut pas mettre en valeur un système de prévision seulement par rapport à des estimations des erreurs de prévision dans



l'avenir. On doit plutôt souligner que les erreurs d'« estimation » dans le passé, telles qu'obtenues par le biais de l'identité (16), ont des distributions inédites qui peuvent différer entre les spécifications  $\rho = 1$  et  $\rho = 2$ , même si leur erreur moyenne est semblable. Il vaut mieux mettre en cause les distributions comparatives de ces erreurs comme, par exemple, l'étendue relative de l'auto-corrélation non modélisée.

### Un prolongement stochastique

L'équation (14), qui rend compte dans le cadre des SVAR des modèles habituellement utilisés pour la projection des effectifs scolaires, se généralise au plan stochastique par le biais de

$$A_1^G = \begin{bmatrix} \phi_{1,1}^1 & 0 & 0 & \cdots & 0 \\ \phi_{2,1}^1 & \phi_{2,2}^1 & 0 & \cdots & 0 \\ 0 & \phi_{3,2}^1 & \phi_{3,3}^1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \phi_{j,j}^1 \end{bmatrix} \quad A_\ell^G = \begin{bmatrix} \phi_{1,1}^\ell & 0 & 0 & \cdots & 0 \\ \phi_{2,1}^\ell & \phi_{2,2}^\ell & 0 & \cdots & 0 \\ 0 & \phi_{3,2}^\ell & \phi_{3,3}^\ell & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \phi_{j,j}^\ell \end{bmatrix} \quad A_p^G = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 \\ \phi_{2,1}^p & 0 & 0 & \cdots & 0 \\ 0 & \phi_{3,2}^p & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix}. \quad (20)$$

Elle est stochastique dans la mesure où les éléments non nuls sont désormais les paramètres qui peuvent être estimés au moyen des données; ils ne sont plus les constantes, les coefficients fixes<sup>7</sup>. Toutefois, le fait qu'il n'y ait pas tant de paramètres à estimer évite de nous heurter aux problèmes d'identification qui menacent les VARS désordonnés. L'équation (14) est généralisée dans le sens où on peut préciser plusieurs modèles de projection par cohorte qui dépassent les suppositions avancées dans le modèle de l'équation (9), laquelle ne fait que susciter les équations matricielles complètement contraintes comme dans l'équation (14).

Afin de comprendre les implications d'une telle spécification pour l'estimation des rapports de survie, il faut réécrire l'équation (9) en tant que moyenne géométrique pondérée

$${}_{j-1}^j \varphi[\rho] = \left[ \prod_{\ell=1}^{\rho} \left( \frac{w_{t-\ell+1,j}^{t-\ell-j+2}}{w_{t-\ell,j-1}^{t-\ell-j+2}} \right) {}_{j-1}^j a^{t-\ell-j+2} \right]^{1/\sum_{\ell=1}^{\rho} ({}_{j-1}^j a^{t-\ell-j+2})} \quad (21)$$

7. Dans la section suivante, on traite la détermination des valeurs à venir pour le kindergarten. En attendant, des paramètres seront laissés dans la première ligne en vue d'un processus autorégressif.

avec les poids  $\alpha_{j-1}^{t-\ell-j+2}$ . Il s'agit typiquement (e. g. Spar 1994; Rushton *et al.*, 1995 : 174-175; Grip, 2004; Grip, 2009) d'une moyenne simple, ainsi

$$\alpha_{j-1}^{t-\ell-j+2} = 1 \text{ pour } \ell = 1, \dots, \rho \quad (22)$$

l'équation (21) se simplifiant en l'équation (9). On peut aussi se servir de ce système de pondération pour spécifier un lissage exponentiel simple,

$$\alpha_{j-1}^{t-\ell-j+2} = \lambda(1 - \lambda)^{\ell-1} \text{ pour } \ell = 1, \dots, \rho. \quad (23)$$

Dans ce cas, le problème de la précision de la durée  $\rho$  se transpose dans celui de la précision du paramètre  $\lambda$  (Montgomery *et al.*, 2008 : 193-206; Grip, 1994 : 9; Siegel, 2002 : 516).

Cependant, si on formule le rapport de progression de l'équation (21),  $\varphi_{j-1}^j$ , en fonction des équations (10), (12), (13), (15) et (20), on aboutit, pour une moyenne en fonction de  $\rho$  cohortes (et pour les classes  $j > 1$ ) à

$$\varphi_{j-1,t}^{j,t+1}[\rho] = \exp\{[(\phi_{j,j-1}^1 - 1)y_{t,j-1}] + \sum_{\ell=1}^{\rho} [\phi_{j,j}^{\ell} y_{t-\ell+1,j} + \phi_{j,j-1}^{\ell+1} y_{t-\ell,j-1}]\}. \quad (24)$$

Il nous faut constater maintenant que le  $\varphi$  s'est doté des termes supplémentaires en  $t$  et  $t-1$  : il s'avère que le rapport de progression pour la transition entre la classe  $j-1$  et la classe  $j$  va dépendre de l'époque. Tout d'abord, s'agissant du terme en  $(\phi_{j,j-1}^1 - 1)y_{t,j-1}$ , pour  $\phi_{j,j-1}^1 \neq 1$ , la croissance de l'effectif vers la classe  $j$  dans l'année  $t+1$  va dépendre de la taille de l'effectif au sein de la même cohorte dans la classe antérieure dans l'année précédente  $y_{t,j-1}$ . Par exemple, pour  $\phi_{j,j-1}^1 < 1$ , il y aura un rajustement montrant que, toutes choses étant égales par ailleurs, autant la taille de la cohorte augmente, autant le rapport de progression utilisé dans les projections va diminuer. En tout cas, ce cadre permet qu'une telle relation puisse être explorée.

Soit qu'il est le résultat des tests, soit qu'on le précise *a priori*, on peut exiger que  $\phi_{j,j-1}^1 = 1$  au moyen de

$$A_1^{G'} = \begin{bmatrix} \phi_{1,1}^1 & 0 & 0 & \dots & 0 \\ 1 & \phi_{2,2}^1 & 0 & \dots & 0 \\ 0 & 1 & \phi_{3,3}^1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \phi_{j,j}^1 \end{bmatrix} \quad (25)$$

au lieu de  $A_1^G$  dans l'équation (20). Mais cela seul ne suffit pas pour que le  $\varphi$  de l'équation (24) puisse se débarrasser des termes supplémentaires en  $t$  et  $t-1$  : cette fois, il s'agit de l'époque des cohortes antérieures qui se prêtent au calcul du moyen rapport pendant les  $\rho$  cohortes antécédentes.

Parce que, même avec  $\phi_{j,j-1}^1 = 1$ , l'équation (24) peut se rédiger de la façon suivante :

$${}_{j-1,t}^{j,t+1}\varphi[\rho] = \exp\left\{\sum_{\ell=1}^{\rho}\phi_{j,j}^{\ell}\left[\left(1 + \frac{\phi_{j,j-1}^{\ell+1}}{\phi_{j,j}^{\ell}}\right)y_{t-\ell,j-1} + y_{t-\ell+1,j} - y_{t-\ell,j-1}\right]\right\} \quad (26)$$

ou, en ce que  ${}_{j-1}^j\varphi^{t-\ell+1} = \exp(y_{t-\ell+1,j} - y_{t-\ell,j-1})$ , comme

$${}_{j-1,t}^{j,t+1}\varphi[\rho] = \exp\left\{\sum_{\ell=1}^{\rho}\phi_{j,j}^{\ell}\left[\left(1 + \frac{\phi_{j,j-1}^{\ell+1}}{\phi_{j,j}^{\ell}}\right)y_{t-\ell,j-1} + \ln {}_{j-1}^j\varphi^{t-\ell+1}\right]\right\} \quad (27)$$

où les  $\{\phi_{j,j}^{\ell}\}$  jouent le rôle des poids  $\{{}_{j-1}^j\alpha^{t-\ell-j+2}\}$  qui se rattachent aux cohortes dans la version multiplicative (donc non logarithmique) de l'équation (21). Toutefois, l'équation (27) se différencie de l'équation (21) d'une autre manière avec le terme  $\left(1 + \frac{\phi_{j,j-1}^{\ell+1}}{\phi_{j,j}^{\ell}}\right)y_{t-\ell,j-1}$ , qui peut, en effet, gonfler ou dégonfler le rapport de progression observé pour une cohorte précédente en fonction de  $y_{t-\ell,j-1}$ . Si on imagine que les rapports  $\{{}_{j-1}^j\varphi^{t-\ell+1}\}$  vont rester fixes, dans le sens où un changement en  $y_{t-\ell,j-1}$  va conduire à un changement proportionnel en  $y_{t-\ell+1,j}$  (parce que le terme  $y_{t-\ell+1,j}$  n'apparaît pas ailleurs dans l'équation [27]), on se retrouve encore une fois avec un modèle de projection par cohorte où les paramètres canoniques vont varier en fonction des tailles relatives des cohortes.

Peut-être est-ce un avantage. Nous l'ignorons à l'heure actuelle car ce n'est pas une perspective qui a figuré dans les modèles de ce type-ci. L'hypothèse se met en valeur dans la rejection des contraintes

$$\phi_{j,j}^{\ell} = -\phi_{j,j-1}^{\ell+1} \quad (28)$$

sur l'équation (20), pour tous les  $j$ . Sinon, l'équation (27) se réduit à

$${}_{j-1}^j\varphi = \exp\left\{\sum_{\ell=1}^{\rho}\phi_{j,j}^{\ell}\ln {}_{j-1}^j\varphi^{t-\ell+1}\right\} \quad (29)$$

qui ressemble à l'équation (21) sauf que, en l'absence d'une contrainte de normalisation supplémentaire, à la manière de

$$\sum_{\ell=1}^{\rho}\phi_{j,j}^{\ell} = 1 \quad (30)$$

on risque que la valeur de  ${}_{j-1}^j\varphi$  ne soit plus une mesure de tendance centrale pour l'ensemble des rapports  $\{{}_{j-1}^j\varphi^{t-\ell+1}\}$ . Que signifierait un tel résultat ? Si, par exemple, pour tous les  $\ell$ ,  $\varphi^{t-\ell+1} > 1$ , et qu'il s'avère que

$\sum_{\ell=1}^p \phi_{j,j}^{\ell} < 1$ , il s'agirait alors d'un rétrécissement de l'estimation de  ${}_{i-1}^j \varphi$  vers l'unité. Mais il reste plusieurs combinaisons que l'on aurait du mal à comprendre sur le plan substantif. En conséquence, la contrainte (30) se recommande en général.

Même avec toutes ces contraintes — les équations (25), (28) et (30) appliquées à l'équation (20) — on ne peut revenir aux équations typiques de la pratique de la projection des effectifs scolaires par cohorte, à l'instar de ceux de l'équation (14). En effet, étant conditionné par le choix de  $\rho$ , on va obtenir les estimations de tous les poids  $\{{}_{i-1}^j \alpha^{t-\ell-j+2}\}$  de l'équation (21), sans avoir besoin de préciser leurs égalités, comme dans l'équation (22), ou s'ils s'atténuent avec le temps, comme dans l'équation (23). Si maintenant on précise la contrainte supplémentaire sur l'équation (20),

$$\phi_{j,j}^{\rho} = 0 \quad (31)$$

pour tous les  $j$ , on teste effectivement l'hypothèse qu'il ne faut se servir que de  $\rho - 1$  rapports (au lieu de  $\rho$ ) pour capter toute l'information dans les séries.

## CONCLUSION

Ces sujets, comme par exemple la durée d'un décalage pour un VAR (Enders, 2004 : 357-358), constituent le gagne-pain des économétriciens. Ce faisant, ceux-ci ont développé de nombreux modèles qui s'avèrent applicables à la prévision démographique par la méthode des composantes. Les rapports canoniques qui relient les comptes ou les taux au sein des cohortes au fil des âges peuvent être exprimés d'une façon matricielle, telle qu'exemplifiée dans ces pages, et isomorphe comme dans le cadre plus vaste de l'économétrie. Aussi, les chercheurs-démographes œuvrant dans le domaine de la prévision gagneraient à faire usage de la théorie et des logiciels existants, car ils préserveraient en même temps la « mémoire » qui est le pilier de la projection démographique par cohorte. De cette manière, nous avons aussi établi une dynamique antérieurement absente dans la prévision effectuée par le biais de l'analyse micro-économétrique des séries chronologiques. Nous avons voulu par là attirer l'attention sur ces possibilités à travers l'exemple des projections des effectifs de population scolaire.

Nous suggérons pour terminer une ébauche de trois pistes prometteuses dans le domaine de la prévision démographique, et qui se dégagent

de nos analyses. La première piste concerne *la stabilité du processus*. On a montré que même les modèles les plus simples dans le cadre de la prévision faisant appel aux rapports de progression se prêtaient à une interprétation en tant que modèle autorégressif pour les séries chronologiques. Preuve a été également faite que ces derniers modèles, en impliquant une structure des erreurs à travers la durée de la série, peuvent servir à estimer à l'avance les erreurs de prévision ; et que l'on peut se servir ultimement de ces données pour rendre stochastiques plusieurs paramètres qui, jusqu'à présent, étaient fixes. Par contre, la tendance au sein de la recherche sur les projections courantes d'effectifs scolaires est de croire que tout change et que « l'histoire de longue durée » de ces effectifs n'a rien à voir avec ce qui s'est passé plus récemment, voire ce qui va se dévoiler dans les années à venir (e. g. Spar 1994 ; Rushton *et al.*, 1995 : 174-175 ; Grip 1994, 2004 ; cf. Campbell, 1997 : 29). Cependant, entre ces deux positions une question demeure sous-jacente : remarque-t-on un processus fixe, immuable et invariable ? Et en quel sens, ou plutôt, de quelle manière ? L'analyse des séries chronologiques est associée aux principes de la stabilité et de la stationnarité, de l'invariance des paramètres et de la dispersion des erreurs. Il existe des tests de ces propositions (e. g. Enders, 2004 : 266-268) ainsi que des transformations assez simples afin que, dans quelques cas, on puisse maintenir ces propositions quand elles ne se tiennent pas par rapport aux séries au niveau de leurs mesures originales. Certes, dans le monde des effectifs scolaires tout change tout le temps ; mais plutôt que de simplement l'assumer sans avoir à se remettre en question, il vaut mieux le démontrer d'une façon formelle.

La seconde piste a trait *aux données exogènes*. Le système des équations (1), (2), (15) (ou [19]), (20) et (6) est un système récuratif. Ainsi, pour chaque classe, tout dépend de l'histoire de cette classe et de la classe antérieure. Nous avons évité jusqu'à maintenant la question fondamentale : d'où viennent les élèves de la première classe, qu'il s'agisse d'un *kinder-garten* ou d'une « vraie » classe 1 ? Si l'on peut faire entrer des élèves venant d'ailleurs, on peut réduire l'erreur qui, autrement, se répandrait partout par le biais d'un modèle complètement endogène et autorégressif du type de celui qui a été proposé ci-dessus lors de l'exploration des intervalles de confiance des prévisions. Il est bien connu qu'il devrait y avoir un rapport entre les naissances qui se produisent cinq, six ou sept années auparavant et les élèves qui arrivent à l'époque  $t$ , mais de quelles naissances s'agit-il, et vers quelle classe ces élèves vont-ils aller ? Il y a beaucoup de migration parmi les jeunes enfants, suivant naturellement les mouvements de leurs mères et pères. Pour les petits secteurs scolaires, en particulier, il est

possible que la variabilité du flux des naissances locales ne se conformera pas assez bien au flux d'élèves, qui est fonction également des autres tendances macro-démographiques (Siegel, 2002 : 56-57). On a déjà constaté que le point d'entrée peut être ambigu : les *kindergartens*, par exemple, ne sont pas toujours obligatoires. Quant à la durée du décalage, il y a les âges « normaux » prescrits par la loi, mais il y a aussi les parents qui retiennent leurs enfants pour une raison ou pour une autre (*e. g.* Grondin *et al.*, 1984), ou, à l'inverse, qui exigent une entrée précoce à l'école (Deschamps et Larivée, 2001). Enfin, la date à laquelle un enfant doit avoir six ans est rarement le premier janvier.

Ainsi, il existe une poignée de facteurs qui profiteraient d'un système omnibus de pondération, entre une ou plusieurs séries de naissances et les deux premières classes, pour plusieurs âges. On peut imaginer, par exemple, deux séries de naissances, l'une pour l'agglomération locale (celle qui correspond au secteur scolaire), l'autre pour un niveau d'agrégation plus élevé. Les deux séries seraient corrélées l'une avec l'autre. Toutefois, la relation entre les deux ne nous intéresse pas, car dans l'optique de la prévision d'effectifs d'élèves, elles sont exogènes. On a déjà rencontré, par le biais de l'équation (7), les modèles VARX (Bierens, 2004) qui se prêtent naturellement à cette tâche. Il s'agit d'augmenter l'équation (1) de la façon suivante :

$$\mathbf{y}_t = \alpha_0 + A_1 \mathbf{y}_{t-1} + \dots + A_\ell \mathbf{y}_{t-\ell} + \dots + A_p \mathbf{y}_{t-p} + B_1 \mathbf{x}_{t-a} + \dots + B_h \mathbf{x}_{t-a-h} + \dots + B_q \mathbf{x}_{t-a-q+1} + \mathbf{u}_t \quad (32)$$

où  $\alpha$  est l'âge minimum pour qu'un enfant puisse commencer l'école,  $1 \leq h \leq q$  est le nombre  $q$  de vecteurs  $\mathbf{x}$  décalés. Si les naissances ne font que s'associer aux effectifs de la première année de l'école primaire, on aura tout de même  $q = 3$  comme minimum<sup>8</sup> ! La longueur  $M$  de ces vecteurs indique le nombre de séries de naissances ; typiquement, on a  $M = 2$  ou même  $M = 1$ , mais en général :

$$\mathbf{x}_{t-a} = \begin{bmatrix} x_{1,t-a} \\ x_{2,t-a} \\ x_{3,t-a} \\ \vdots \\ x_{M,t-a} \end{bmatrix} \quad \mathbf{x}_{t-a-h} = \begin{bmatrix} x_{1,t-a-h} \\ x_{2,t-a-h} \\ x_{3,t-a-h} \\ \vdots \\ x_{M,t-a-h} \end{bmatrix} \quad \mathbf{x}_{t-a-q+1} = \begin{bmatrix} x_{1,t-a-q+1} \\ x_{2,t-a-q+1} \\ x_{3,t-a-q+1} \\ \vdots \\ x_{M,t-a-q+1} \end{bmatrix}. \quad (33)$$

8. Dans l'année  $t$ , les enfants qui atteignent l'âge  $\alpha$  sont nés au cours des années  $t - \alpha$  et  $t - \alpha - 1$ . Par conséquent, il faut estimer deux rapports de progression de la naissance à l'école. Pour chacun de ces rapports, il y a un décalage de deux  $q$  (exprimant la longueur du décalage) pour former le rapport de progression dans la cohorte précédente. Mais étant donné qu'ils se chevauchent dans l'année  $t - \alpha - 1$ , on obtient trois  $q$  comme minimum.

Ces vecteurs se relient aux effectifs scolaires par l'intermédiaire de

$$B_1 = \begin{bmatrix} \tau_{1,1}^1 & \tau_{1,2}^1 & \tau_{1,3}^1 & \cdots & \tau_{1,M}^1 \\ \tau_{2,1}^1 & \tau_{2,2}^1 & \tau_{2,3}^1 & \cdots & \tau_{2,M}^1 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix}_{J \times M} \quad B_h = \begin{bmatrix} \tau_{1,1}^h & \tau_{1,2}^h & \tau_{1,3}^h & \cdots & \tau_{1,M}^h \\ \tau_{2,1}^h & \tau_{2,2}^h & \tau_{2,3}^h & \cdots & \tau_{2,M}^h \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix}_{J \times M} \quad B_q = \begin{bmatrix} \tau_{1,1}^q & \tau_{1,2}^q & \tau_{1,3}^q & \cdots & \tau_{1,M}^q \\ \tau_{2,1}^q & \tau_{2,2}^q & \tau_{2,3}^q & \cdots & \tau_{2,M}^q \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix}_{J \times M} \quad (34)$$

La dernière piste s'inscrit dans *une perspective bayésienne*. Le cadre matriciel pour la prévision des séries temporelles s'affilie assez naturellement à une perspective bayésienne (Montgomery *et al.*, 2008 : 354). Les modèles comme ceux de l'équation (14) sur lesquels nous avons mis l'accent peuvent être généralisés (voir l'équation [20]) afin de permettre une plus grande concordance avec les données et, peut-être, créer des prévisions supérieures. Les macro-économistes sont parvenus à améliorer les prévisions des modèles VAR sans contraintes en les mêlant aux modèles plus simples, par exemple celui de la marche aléatoire, qui précisent les matrices  $A$  beaucoup plus structurés (*e. g.* Del Negro et Schorfheide, 2003). En comparaison, on pourrait dire d'un point de vue bayésien que les méthodes démographiques habituelles, exprimées notamment au travers des équations (14) et (15), relèvent d'une croyance tellement forte relativement à leur antériorité (un « prior ») que nulle évidence ne va les ébranler. Pourtant, on peut imaginer quelques estimateurs qui, toujours dans l'esprit de Bayes, commenceraient avec les estimations d'un modèle comprenant plusieurs contraintes relaxées, comme dans l'équation (20), et qui rétréciraient l'information vers le niveau des paramètres impliqués par les modèles démographiques canoniques. Mais ce type de proposition nécessiterait davantage d'observations et de réflexions conséquentes dans le cadre de recherches ultérieures.

Nous arrivons au terme de cette étude sans avoir épuisé toutes les idées suscitées par les modèles « simples » pour la projection par cohorte d'effectifs scolaires, tandis qu'il reste bien d'autres problèmes liés à la prévision démographique. Dans ce dernier registre, les méthodes de Lee et Carter (1992) pour la prévision de taux de mortalité par cohorte suggèrent l'élaboration d'une structure qui pourrait être également abordée à partir des idées et des identités esquissées dans ces pages.

## BIBLIOGRAPHIE

- ALHO, J. M., et B. D. SPENCER. 2005. *Statistical demography and forecasting*. New York, Springer, 410 p.
- BIERENS, H. J. 2004. « VAR models with exogenous variables », Pennsylvania State University, Department of Economics. *Site de l'auteur*, [En ligne], [http://econ.la.psu.edu/~hbierens/EasyRegTours/VAR\\_Tourfiles/VARX.PDF](http://econ.la.psu.edu/~hbierens/EasyRegTours/VAR_Tourfiles/VARX.PDF) (consulté le 20 décembre 2008).
- BOOTH, H. 2006. « Demographic forecasting : 1980 to 2005 in review », *International Journal of Forecasting*, 22, 3 : 547-581.
- BUTZ, W. P., et M. P. WARD. 1979. « The emergence of countercyclical American fertility », *American Economic Review*, 69, 3 : 318-328.
- CAMPBELL Jr., H. S. 1997. « School enrollment in an extended demographic model », *Journal of Economic & Social Measurement*, 23, 1 : 27-42.
- DEL NEGRO, M., et F. SCHORFHEIDE. 2003. « Take your model bowling : Forecasting with general equilibrium models », *Economic Review*, 88, 4 : 35-50.
- DEMOGRAPHICS AND WORKFORCE. Sans date. « Fall membership projection methodology », Université de Virginie, Weldon Cooper Center for Public Service. [http://www.coopercenter.org/demographics/sitefiles/documents/pdfs/schoolforecasts/proj\\_methodology.pdf](http://www.coopercenter.org/demographics/sitefiles/documents/pdfs/schoolforecasts/proj_methodology.pdf) (consulté le 13 juillet 2009).
- DESCHAMPS, I., et S. LARIVÉE. 2001. « L'impact développemental de la dérogation à l'âge d'admission au primaire », *Revue des sciences de l'éducation*, 27, 3 : 671-686.
- DESNOYERS, Y. 2001. *L'effet de la richesse sur la consommation aux États-Unis*. Document de travail 2001-14, Banque du Canada.
- DUCHÊNE, J. 1999. « Le traitement de l'incertitude dans les projections de population », *Quaderns d'Estadística i Investigació Operativa (Questiio)*, 23, 1 : 113-128.
- ENDERS, W. 2004. *Applied econometric time series* (deuxième édition). New York, John Wiley & Sons Inc., 460 p.
- FAVERO, C. A. 2007. « Model evaluation in macroeconometrics : From early empirical macroeconomic models to DSGE models », Université Bocconi, IGIER, *Working Papers*. [http://didattica.unibocconi.it/mypage/upload/48917\\_2008-1006\\_041428\\_favero\\_oxdfh.pdf](http://didattica.unibocconi.it/mypage/upload/48917_2008-1006_041428_favero_oxdfh.pdf) (consulté le 18 décembre 2008; dernier changement le 12 mai 2008).
- GOUVERNEMENT DU QUÉBEC. 2009. *Projet de règles budgétaires pour l'année scolaire 2009-2010*. Commissions scolaires, document complémentaire, ministère de l'Éducation, du Loisir et du Sport. <http://www.rtsq.qc.ca/apds/accueil/documents/Compl2009-2010.pdf> (consulté le 15 août 2009).
- GRANGER, C. W. J. 1969. « Investigating causal relations by econometric models and cross-spectral methods », *Econometrica*, 37, 3 : 424-438.



- GRANGER, C. W. J. 2007. « Forecasting—looking back and forward : Paper to celebrate the 50th anniversary of the Econometrics Institute at the Erasmus University, Rotterdam », *Journal of Econometrics*, 138, 1 : 3-13.
- GRIP, R. S. 2004. « Projecting enrollment in rural schools : A study of three Vermont school districts », *Journal of Research in Rural Education*, 19, 3 : 1-6.
- GRIP, R. S. 2009. « Does projecting enrollments by race produce more accurate results in New Jersey school districts ? », *Population Research and Policy Review*. <http://www.springerlink.com/content/m6932k1m9x734542/> [version du 13 février 2009].
- GRONDIN, S., P. DESHAIES et L. P. NAULT. 1984. « Trimestres de naissance et participation au hockey et au volleyball », *La Revue québécoise de l'activité physique*, 2 : 97-103.
- GUJARATI, D. N. 2004. *Économétrie*. B. Bernier, trad. Bruxelles, Éditions De Boeck, 1024 p.
- INSERM (Expertise collective). 2003. *Tabagisme : Prise en charge chez les étudiants*. Paris, Les Éditions de l'INSERM.
- JACOBY, E. G. 1959. *L'application des méthodes de projection aux effectifs scolaires*. Paris, UNESCO, Études et documents d'éducation, n° 32, 48 p.
- LAMY, M. 2004. « Perspectives scolaires : des perspectives dérivées à l'application de la méthode des flux », dans G. CASELLI, J. VALLIN et G. WUNSCH, éd. *Démographie : analyse et synthèse*, V, *Histoire du peuplement et prévisions*. Paris, INED : 277-284.
- LE BRAS, H. 1987. « Nature et limites des prévisions de population », dans INED, éd. *Les projections démographiques : Actes du VIII<sup>e</sup> colloque national de démographie*. Grenoble, mai 1987, Paris, PUF : 13-39.
- LE BRAS, H. 2005. *La démographie*. Paris, Éditions Odile Jacob, 455 p.
- LEDENT, J. 1995. « Vers une projection des familles selon leurs caractéristiques principales », *Cahiers québécois de démographie*, 24, 1 : 3-33.
- LEE, R. D., et L. R. CARTER, 1992. « Modeling and forecasting U. S. mortality », *Journal of the American Statistical Association*, 87, 419 : 659-671.
- LEE, R. D. et S. TULJAPURKAR, 1994. « Stochastic population forecasts for the United States : Beyond high, medium, and low », *Journal of the American Statistical Association*, 89, 428 : 1175-1189.
- LÉRIDON, H. 2003. « Certitudes et incertitudes de la prévision démographique », *C. R. Géoscience*, 335 : 627-635.
- MCDONALD, J. 1980. « Birth time series models and structural perspectives », *Journal of the American Statistical Association*, 75, 369 : 39-41.
- MONTGOMERY, D. C., C. L. JENNINGS et M. KUHLACI. 2008. *Introduction to time series analysis and forecasting*. New York, John Wiley & Sons, Inc., 445 p.
- MORRISON, P. A. 2000. « Forecasting enrollments for immigrant entry-port school districts », *Demography*, 37, 4 : 499-510.

- NAMBOODIRI, K. 1991. *Demographic analysis : A stochastic approach*. San Diego, Academic Press Inc., 370 p.
- PROJET SUR LE DÉSÉQUILIBRE DES RESSOURCES GOUVERNEMENTALES DANS LA MRC DE THÉRÈSE-DE-BLAINVILLE. 2007. « Le déséquilibre freine notre développement : l'éducation ... est au ralenti », Dossier de presse, [En ligne], [http : //www.mrctheresedeblainville.qc.ca/desequilibre/pdf/d\\_presse\\_education.pdf](http://www.mrctheresedeblainville.qc.ca/desequilibre/pdf/d_presse_education.pdf) (consulté le 15 août 2009).
- RENSHAW, A., et S. HABERMAN, 2003. « Lee–Carter mortality forecasting : a parallel generalized linear modelling approach for England and Wales mortality projections », *Journal of the Royal Statistical Society : Series C (Applied Statistics)*, 52, 1 : 119-137.
- RUSHTON, G, M. P. ARMSTRONG et P. LOLONIS. 1995. « Small area student enrollment projections based on a modifiable spatial filter », *Socio-Economic Planning Sciences*, 29, 3 : 169-185.
- SIEGEL, J. S. 2002. *Applied demography : Applications to business, government, law, and public policy*. San Diego, Academic Press Inc., 686 p.
- SIMS, C. A. 1980. « Macroeconomics and reality », *Econometrica*, 48, 1 : 1-48.
- SIMS, C. A. 1986. « Are forecasting models usable for policy analysis ? », *Federal Reserve Bank of Minneapolis, Quarterly Review*, 10, 1 : 2-16.
- SPAR, M. 1994. « An evaluation of techniques for projecting grade-progression ratios », *Educational Research Quarterly*, 18, 2 : 5-22.
- STATE OF CALIFORNIA. 2008. « California Public K-12 Graded Enrollment and High School Graduate Projections by County, 2008 Serie », Department of Finance, *Demographic, Economic, and Financial Research*. [http : //www.dof.ca.gov/research/demographic/reports/projections/k-12/](http://www.dof.ca.gov/research/demographic/reports/projections/k-12/) (consulté le 10 janvier 2009).
- STOCK, J. H., et M. W. WATSON. 2001. « Vector autoregressions », *Journal of Economic Perspectives*, 15, 4 : 101-114.
- TULJAPURKAR, S. 1992. « Stochastic population forecasts and their uses », *International Journal of Forecasting*, 8, 3 : 385-391.
- UNESCO. 2009. « Taux de transition », Institut de statistique de l'UNESCO, Montréal, Québec. [http : //www.uis.unesco.org/i\\_pagesfr/indspec/tecspe\\_tran.htm](http://www.uis.unesco.org/i_pagesfr/indspec/tecspe_tran.htm) (consulté le 10 juillet 2009).
- WACHTER, K. W. 1992. « Variabilité aléatoire des phénomènes démographiques : enseignements des séries paroissiales de Wrigley et Schofield », dans A. BLUM, N. BONNEUIL et D. BLANCHET, éd. *Modèles de la démographie historique*. Paris, INED : 77-98.
- WANG, H. 2008. *Essays on smoking and mortality in the United States*. Doctoral thesis submitted to the Graduate Group in Demography, University of Pennsylvania, Philadelphia.
- WANG, H., et S. H. PRESTON, 2009. « Forecasting United States mortality using cohort smoking histories », *Proceedings of the National Academy of Sciences*, 106, 2 : 393-398.

## ABSTRACT

**Herbert L. SMITH**

### **Application of Time-Series Analysis to the Projection of School Enrollments by Cohort**

This article shows that we can re-write several demographic models for cohort projections as transpositions of the econometric vector auto-regression (VAR) model. In so doing, we give the method of cohort projection a stochastic framework that extends its applicability. This is demonstrated via an example involving the projection of school enrollments. We emphasize a series of equations that allow us to check the validity of several modeling choices that are otherwise made on the basis of habit alone.